

The Challenges of Researching Algorithms

In the debate on algorithmic accountability, and platform responsibility more specifically, the contribution of the social researcher is immense. In this set of posts, researchers reflect upon broad themes of control and agency — not only that which is faced by the data subject, but also by the researcher who relies on proprietary platforms to understand how these systems operate and interact with users. This research bears relevance to policy debates, because it provides evidence of ways in which automated systems shape consumer and citizens and look beyond conventional recommendations of transparency or openness.

Carolyn Gerlitz – Negotiating methodological bias

Mark Coté – Algorithmic Power and Art in the Age of Digital Reproduction

Nathaniel Tkacz – The Bounded Rationality of Algorithmic Accountability



Negotiating methodological bias

In the debate on algorithmic power and accountability, Carolyn Gerlitz, Assistant Professor in New Media and Digital Culture, argues that sociological and media studies research on online platforms (e.g. Twitter or Facebook) should aim for critical reflexivity, acknowledging and being transparent about biases inscribed not only in data or the technicity of platforms, but also the analytical tools we use.

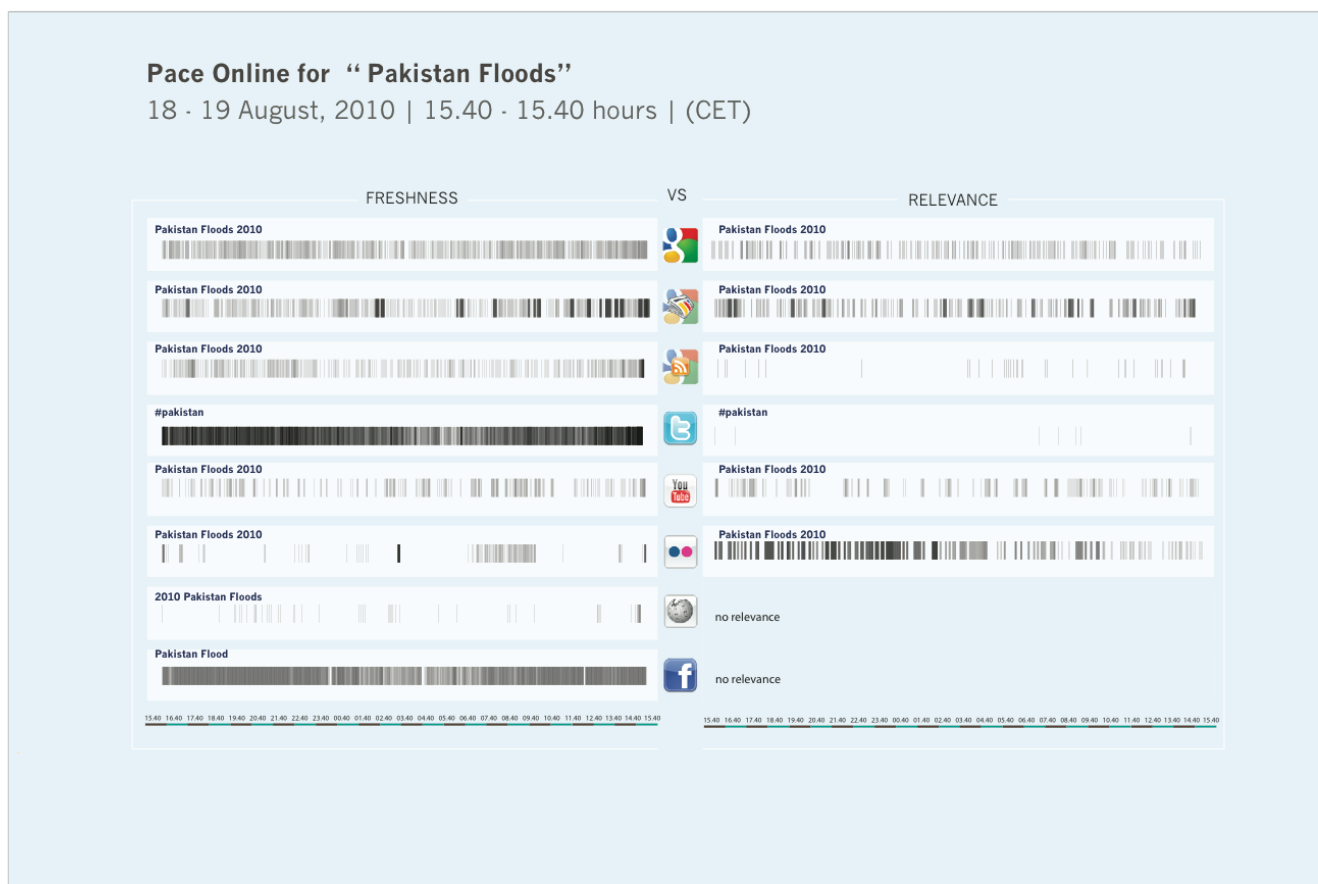
Doing digital research presents researchers with a challenge: online media, their data, algorithms and tools for data extraction, analysis and visualisation are designed to be used in multiple contexts. Their purpose is not fixed – they can be deployed for commercial audience research, intelligence collection, issue analysis or social research amongst others. However, despite being multi-valent objects – that is, speaking to different value registers – data and tools may still come with analytical inscriptions (the ways in which media shape their data) or biases (Marres & Gerlitz 2015) which prioritise some forms of research over others. Some of these biases are fairly obvious, such as the focus on positive affect on Facebook (through the presence of Liking), whilst others are more ambiguous, like the pull towards proportional measures on Twitter (facilitated through its focus on trending topics).

Collecting data from platforms, repurposing their algorithmic sorting capacities, deploying platform APIs, scrapers or free online tools for analysis may thus result in “cascades of inscriptions” or biases (Ruppert et al. 2013), as each element of the methodological apparatus comes with its specific, but not necessarily aligning bias. Doing digital research thus poses the question of how these inscriptions a) play out together and b) align or mal-align with the researchers’ objective. In recent work with Noortje Marres (2015), we claim that rather than trying to eliminate any misfit between the methodological bias of data or tools and one’s own research objective, it is worth working with these interface effects and embracing mal-alignment. This could allow for a reflexive stance in situations when the medium is both the site of critical inquiry and also key element of the research apparatus. By doing so, researchers learn more about their own analytical propositions, how media constrain or advance them, and invites experimentation with methodologies without overcorrecting certain types of bias, a process which can pose its own problems.

Let me illustrate my proposal for inviting mal-alignment with by drawing on a previous study on how different devices, including search engines and platforms, organise content in real-time

(Weltevrede et al. 2014). Rather than trying to make their underlying algorithmic and calculative practices transparent, we instead built a software tool that allowed us to trace how the algorithms operate in action. Focusing on both fresh (that is, most recent), and recommended (that is, algorithmically selected), content, we adjusted the tool to each platform, issued a query every five minutes and recorded the number of new results, which are visualised as lines in the chart below. The resulting bars visualise the intensity and rhythm at which the devices deliver new content for the chosen query.

We found continuous pace on Twitter and Google, bulky pace on batch upload platforms like Flickr and stale pace with little updates on Google Blog Search. The pace and temporality of each platform or engine, the project concluded, is dependent on the interaction of the algorithm with user activities (content production), the query and the time. The alignment between method and medium was not entirely neat as the categories of fresh and recommended content do not fit all platforms. But it allowed us to explore algorithms operating in situated contexts and creating the specific temporalities of each platform or engine.



As seen from this example, exploring mal-alignment offers a different perspective on two key concerns for working with biased or opaque data and tools. Firstly, it offers an alternative to the desire for complete transparency, for instance by reading code or reconstructing algorithmic ordering. Instead, embracing mal-alignment allows researchers to study the effects and biases of the medium ‘in action’, which can be understood as specific and situated effects of media operating in methodological assemblies. Secondly, there are implications for our understanding of medium-specificity, which cannot be considered as a discrete property of data, tools or media, but instead should be regarded as a situated and distributed, as seen in the assembly of real-time streams, which follow their distinctive rhythms. The specificity of a medium does not solely arise from itself, but from its operations in different contexts.

The focus on methodological alignment and situated observation of algorithms thus offers an alternative to the desires for transparency. Tools, data and media do not exist in isolation and in order to account for their methodological biases, one needs to perceive this bias as distributed accomplishment.



Algorithmic Power and Art in the Age of Digital Reproduction

Mark Coté, Lecturer in Digital Culture & Society at King's College London and AHRC-funded researcher on Big Social Data, considers how society might exploit technological advances in algorithmic systems and data for uses in other fields, such as art and politics.

The algorithmic power expressed in complex assemblages of datafication has created highly asymmetrical conditions where we ubiquitously generate data over which we have little agency. More robust regulatory oversight, **as suggested by Frank Pasquale** is one path to articulating a more equitable data public sphere. This data deluge also entails normative pitfalls of a promised 'N=all' verisimilitude. Carolin Gerlitz's **suggestions** that researchers regard data in their medium specificity, and thus foreground cascading biases as an opportunity for critical reflection, is a welcome and innovative approach to new empirical methods using social big data (**SoBigData**). But what of the lived experience of SoBigData, our ubiquitous and habituated collective data generation? This provocation suggests that the SoBigData underpinning algorithmic power can equally be regarded as a collective cultural resource for critical and creative use increasing our data agency.

I make this proposal in the light of a series of interdisciplinary research projects I've worked on which explore the techno-cultural conditions of the human-nonhuman relations comprising these complex assemblages of datafication. These Arts and Humanities Research Council (AHRC) and **European-funded projects** respectively explore SoBigData in terms of **access**; questioning what of that data is, could or should be *open*; how we might discern **cultural** dimensions in data that otherwise drives neoliberal drives of managerial efficiencies and productivity; and the **creative potential** in the data we collectively generate. All of these projects are fundamentally interdisciplinary with a non-prescriptive focus on increasing agency across processes of datafication.

Here I propose that fresh insights can be gained by revisiting Walter Benjamin's 'The Work of Art in the Age of Mechanical Reproduction'. This classic work, highly influential in media theory and cultural studies, is usually deployed for understanding the cultural, aesthetic, and potentially revolutionary dimensions of mechanisation in late modernity. But there are basic formulations in Benjamin's critique that I read as apposite for understanding the algorithmic power of SoBigData.

First, Benjamin focused on the mechanical reproduction of art in order to better understand the new temporalities of industrial mass production. Algorithmic power, driven by machine learning, deep learning and neural networks, is reconfiguring production and consumption under an accelerated temporality exceeding the parameters of human cognition. Thus while the mass industrialisation of film expresses the repetition of Fordist production, the predictive analytics articulating the cultural flows and circuits of production reveal our deeply constitutive role in the social data factory.

Second, consider Benjamin's assertion that mechanical reproduction, as exemplified in film, "enriched our field of perception" (235). What he emphasised is how film opens life to analysis on a radically expanded spectrum and with a precision heretofore unknown. What was true for film is even more so for datafication. Our ubiquitous generation of data opens quotidian life to analysis and action at previously unknown levels of fine granulation. Benjamin was indicating how technical reproduction recalibrates the process of apperception; that is, how we apprehend an object in relation to the self, or, how we can perceive and think about ourselves and the world around us. Apperception is thus a mediated process that transpires on a general and specific level. The former is transcendental apperception, which is the abstract conditions of the Cartesian cogito. In short, how does SoBigData change the ontological parameters of *I think* and thus the lived condition of the data human? The second is empirical apperception, that which allows us to recognise a particular object. Specifically, this is what allows us to distinguish the phenomenon of

the object from its representation. This brings us to the core problematic of data-driven methodologies and new kinds of empirical research. This may further the insights suggested by Gerlitz and enrich a deeper understanding across the data public sphere.

Finally, Benjamin targeted the political and aesthetic possibilities inherent in “the mutual penetration of art and science” (236) which underpinned mechanical reproduction. Benjamin remains a potent conceptual force because he evinced the techno-cultural foundations of politics and art. Can we not, thus, also pursue a creative diversion of algorithmic power to something more than the mere reproduction of the data human from circuits of consumption, surveillance and productivity? The **Persona Non Data project** on which I collaborated with the artists Salvatore Iaconesi and Oriana Persico is but one of endless possible creative openings of access to and use of SoBigData to cultivate new modes of agency in our lived processes of datafication. Benjamin sought new cultural and aesthetic practices in the age of the industrial factory “useful for the formulation of revolutionary demands in the politics of art” and to short-circuit the “processing of data in the Fascist sense” (218). What held for mechanical reproduction seems even more urgent for the algorithmic power driving our social data factory.



The Bounded Rationality of Algorithmic Accountability

Nathaniel Tkacz, Associate Professor at the Centre for Interdisciplinary Methodologies at the University of Warwick, writes that any study of algorithmic accountability can not be disassociated from the wider context of how humans make decisions.

From the archives of mid-20th Century managerial thought:

What is our mental image of the decision maker? Is he a brooding man on horseback who suddenly rouses himself from thought and issues an order to a subordinate? Is he a happy-go-lucky fellow, a coin poised on his thumbnail, ready to risk his action on the toss? Is he an alert, gray-haired businessman, sitting at the board of directors' table with his associates, caught at the moment of saying 'aye' or 'nay'? Is he a bespectacled gentleman, bent over a docket of papers, his pen hovering over the line marked (X)?

All of these images have a significant point in common. In them, the decision maker is a man at the moment of choice, ready to plant his foot on one or another of the routes that lead from the crossroads. All the images falsify decision by focusing on its final moment. All of them ignore the whole lengthy, complex process of alerting, exploring, and analysing that precede the final moment.

Reflecting the managerial ideals of its time, this colourful and mildly problematic passage is taken from the opening page of Hebert Simon's *The New Science of Management Decision*, first published in 1960. Simon was one of those rare figures to leave a mark on a number of fields, from computer science, psychology and economics, to public administration and management. In 1978 he was awarded the **Nobel Prize** in Economic Sciences “for his pioneering research into the decision-making process within economic organisations”. Today Simon is perhaps best remembered for his related concepts of ‘bounded rationality’ and ‘satisficing’ which undermined the idea of a rational economic actor at the heart of neoclassical economics and helped paved the way for its behavioural counterpart.

On display in the opening passage are the beginnings of Simon's *procedural* approach to decision making. Decision is not to be confused with some kind of spontaneous fiat moment. Decisions are




complex, tough; they require any number of inputs and unfold over time. Such processes could be studied, mapped out, improved – a new science of management decision is born.

Besides this focus on process, the new science was founded on a peculiar conflation: “What part does decision making play in managing? I shall find it convenient to take mild liberties with the English language by using ‘decision making’ as though it were synonymous with ‘managing’”. The science of “management decision” becomes management as decision. What does this management-as-decision-as-process look like? Simon suggested all decisions pass through three phases: *intelligence*, or “searching the environment for conditions calling for decision”; *design*, or “inventing, developing, and analysing possible courses of action”; and *choice*, where a selection between alternatives is made.

Now understood as processes comprising intelligence, design and choice phases, Simon claimed that all decisions exist on a continuum of *programmability*: “Decisions are programmed to the extent that they are repetitive and routine, to the extent that a definite procedure has been worked out for handling them so that they don’t have to be treated *de novo* each time they occur”. Decisions that are “nonprogrammed” are “novel, unstructured, and consequential” with “no cut-and-dried method for handling the problem”. The notion of “programme” was directly borrowed from “the computer trade”, as a “detailed prescription or strategy that governs a sequence of responses of a system to a complex task environment”. Importantly, though, this was “an *adaptive* response of the system to the situation” (my emphasis). A programme is not solely a set of rules or formal procedures, as belonging to a Weberian bureaucracy, for example, but rather contains a number of strategic responses to the contingencies of organisational life and beyond. A programme incorporates Weber’s “body of rules” but such that they can now be operationalised without the direct oversight by an “official”. This ‘adaptive’ capacity of rules is roughly akin to **Andrew Goffey’s description of algorithms** as “machinic discourse”.

For Simon, the now-obvious next step was to point out that programmable decisions could and were already being automated, i.e. turned into actual programmes. Lower and middle management decisions were being replaced by “large-scale data processing” and the “tools of operations research”. The apotheosis of Simon’s vision: “The automated factory of the future will operate on the basis of programmed decisions produced in the automated office beside it”. And while non-programmable decisions were by definition unable to be automated, they too would “soon undergo as fundamental a revolution as the one which is currently transforming programmed decision making”.

From the two ends of the programmability continuum come two trajectories. The automation of decision and therefore management leads into artificial intelligence research. We can see traces of it in contemporary concerns over the power of algorithms to make decisions. The second, not automated but still revolutionarily transformed practice of decision making, leads into the field of Decision Support Systems (DSS). It is a field founded directly on Simon’s ideas and persists in updated form in today’s business intelligence and analytics industry. Today we speak of *data-driven* decision making for decisions that resist full automation. The trajectories are, of course, the converse of one another. They are two articulations of management-as-decision-as-process. The threshold from one to the other marks the domain (and limits) of automation from that of the uniquely human decision maker. This line or threshold now climbs the corporate ladder like the rest of us.

On 26 January 2016 Professor Frank Pasquale gave a **public lecture at the LSE** on ‘The Promise (and Threat) of Algorithmic Accountability’. His talk covered a number of concerns around algorithms and the emerging data paradigm, including how we often don’t know where the data that feeds algorithms comes from or what it includes; that there are possible biases in algorithmic models; and that there is a general ‘opacity’ surrounding the deployment of algorithms. Pasquale’s concerns are reasonable, as is the wider critique he develops in his recent book, ***The Black Box Society***. However, is the call for ‘algorithmic accountability’ underpinned by a wider desire to make society more ‘intelligible’ the best we can hope for? A new *will to account*? Besides the obvious fact that transparency is a political dead end, any critical engagement with algorithms must  move

beyond the desire to see, sort, rank or moralise. We must move beyond a Spaghetti Western criticism, where the end game is the sorting of algorithms into The Good, The Bad, and The Ugly. Such processes are better left to the algorithms. Instead we might ask how we arrived at a situation where holding these technical actors to account became a pressing issue? What kind of power-thing is an algorithm? What kinds of relations, strategies, confections, limits, and so on, are carried along with it? It turns out there was nothing mild about the liberties taken by Herbert Simon when he turned management into a decision procedure susceptible to automation.

This blog gives the views of the authors and does not represent the position of the LSE Media Policy Project blog, nor of the London School of Economics and Political Science.

This post was published to coincide with a workshop held in January 2016 by the Media Policy Project, 'Algorithmic Power and Accountability in Black Box Platforms'. This was the second of a series of workshops organised throughout 2015 and 2016 by the Media Policy Project as part of a grant from the LSE's Higher Education Innovation Fund (HEIF5). To read a summary of the workshop, please click [here](#).

April 22nd, 2016 | [Algorithmic Accountability, Featured](#) | [0 Comments](#)

5

